THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 52, No. 8, October, 1973
Printed in U.S.A.

Optimum Mean-Square Decision Feedback Equalization

By J. SALZ

(Manuscript received April 20, 1973)

In this work we report new results relating to decision feedback equalization. The equalizer and the transmitting filter are optimized in a PAM data communication system operating over a linear noisy channel. We use a mean-square error criterion and impose an average power constraint at the transmitter. Assuming correct past decisions, an explicit formula for the minimum attainable mean-square error is given. The possible advantages of signaling faster than the Nyquist rate while decreasing the number of levels to maintain the same information rate are investigated. It is shown that, in all cases of practical interest, signaling faster than the Nyquist rate, while keeping fixed the information rate, increases the mean-square error. Finally, to illustrate the use of the results, application is made to a cable channel where the loss in dB varies as the square root of frequency. Various asymptotic formulas and curves are provided to exhibit the relationships between the quantities of interest.

I. INTRODUCTION

A great deal of research, particularly in the past decade, has been expended on the problem of linear equalization. This has yielded a considerable body of theory and technology making possible the design of apparatus for successfully combating intersymbol interference in PAM data transmission systems operating over noisy linear channels where delay distortion predominates. Since linear equalizers must compensate for the channel characteristics in the presence of noise, they cannot be expected to perform well over severely frequency-attenuating channels or channels possessing nulls in the amplitude characteristic.

Interest in the high data rates over voiceband and cable channels inevitably leads to the search for more effective equalization methods. Faster pulse rates place signal energy well within the badly attenuated portion of the transmission spectrum, resulting in severe intersymbol interference correctable by linear methods only at the expense of a significant enhancement of the noise.

A "bootstrap" technique, commonly referred to as "decision feedback," when combined with linear equalization can yield significant performance improvement.[1,2] In this method the samples of the pulse tails (postcursors) interfering with subsequent or future data symbols are subtracted without incurring a significant noise penalty. The effect of pulse tails (precursors) which occur prior to detection and interfere with past symbols is minimized by a conventional linear equalizer.

Much has been written about this subject. In a fundamental paper where an excellent bibliography can be found, Robert Price[3] demonstrated quantitatively the merits of decision feedback equalization in certain applications.

In this work we jointly optimize the receiving and transmitting filters in a PAM data transmission system employing decision feedback. The chief difference between our work and Price's is in the choice of performance criterion. We minimize mean-square error while Price maximized the signal-to-noise ratio under the constraint that the overall intersymbol interference be zero. Our criterion is not as stringent as Price's and allows trade-offs between added noise and intersymbol interference. Monsen[4] also investigated some aspects of our problem but did not arrive at a complete solution. Our chief contribution is an explicit formula for the minimum mean-square error (MSE). The simplicity of the formula makes possible detailed investigation of optimized system performance.

In Section II the model is stated and the problem is formulated. In Section III the receiving filter is optimized and in Section IV the transmitter filter is optimized. In Section V we examine the problem of signaling faster than the Nyquist rate and finally in Section VI we use our results to investigate in detail the performance of a data system operating over a cable channel.

II. THE MODEL AND PROBLEM FORMULATION

The system model under investigation is depicted in Fig. 1. The data signal denoted by $D(t)$ is passed through the transmitting filter having an impulse response $s(t)$ and giving rise to an average transmitted power P. The data symbols $\{a_n\}_{-\infty}^{\infty}$ are independently picked at the rate $1/T$ and take on values with equal probability from the set $\{\pm 1, \pm 3 \pm 5 \cdots \pm (L-1)\}$ where L is an even integer. The resulting signal is admitted to a linear channel characterized by an impulse response $h(t)$. The received signal plus noise is processed by the equalizer which is comprised of a linear filter having impulse

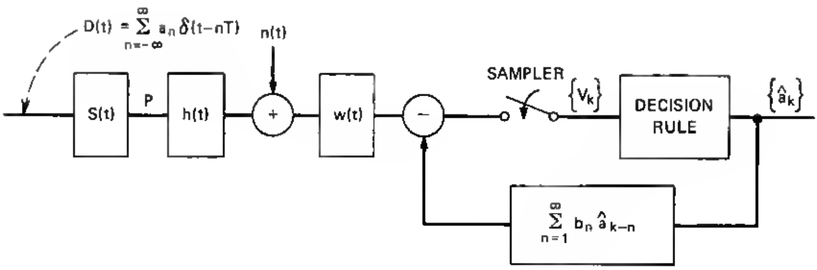

Fig. 1—Block diagram of the model.

response $w(t)$, a sampler, a decision rule, and a feedback digital filter characterized by the infinite set of real numbers $\{b_n\}_1^\infty$. The added noise $n(t)$ is a zero-mean white random process with double-sided spectral density $N_0/2$. The output data symbols are denoted by $\hat{a}_n$.

The general problem we would like to solve is the minimization of

$$\text{MSE} = E\{v_k - \hat{a}_k\}^2$$

with respect to the set of square integrable functions $\{s(t), w(t)\}$ and the infinite sequence of numbers $\{b_n\}$. This is to be carried out when the channel impulse response $h(t)$, transmitted power P, and a decision rule are given. The symbol $E(\cdot)$ denotes expectation with respect to all the random variables.

The nonlinear relation between the estimated symbols $\{\hat{a}_n\}$ and the input symbols $\{a_n\}$ makes this problem mathematically intractable. However, by assuming that past decisions have been correct, we can begin to approach the problem. The resulting MSE must then be interpreted as a lower bound on the true MSE. Alternatively, if no errors have occurred in the past, the MSE under this assumption provides an indication of the noise immunity of the system (including residual intersymbol interference).

Let $r(t) = s(t)*h(t)*w(t)$ denote the overall impulse response, where $*$ denotes convolution. Under the assumption of correct past decisions, the received sample taken at time $t = kT$ is

$$v_k = \sum_{n=-\infty}^{\infty} r_n a_{k-n} - \sum_{n=1}^{\infty} b_n a_{k-n} + n(t)*w(t)\big|_{t=kT}$$

and the mean-square error is then by definition

$$\text{MSE} = E\Bigg\{\sum_{n=-\infty}^{-1} r_n a_{k-n} + \sum_{n=1}^{\infty} (r_n - b_n) a_{k-n} + n(t)*w(t)\big|_{t=kT} + (r_0 - 1) a_k\Bigg\}^2.$$

A straightforward calculation gives

$$\text{MSE} = \sigma_a^2 \sum_{n=-\infty}^{-1} r_n^2 + \sigma_a^2(r_0 - 1)^2 + \sigma_a^2 \sum_{n=1}^{\infty} (r_n - b_n)^2 + \sigma^2,$$

where

$$\sigma_a^2 = E\{a_n\}^2 = \frac{L^2 - 1}{3}$$

and

$$\sigma^2 = \frac{N_0}{2} \int_{-\infty}^{\infty} w^2(t)dt.$$

It can be immediately concluded that the mean-square error is minimized by setting $b_n = r_n$, $n = 1, 2, \cdots$, which eliminates the feedback coefficients $\{b_n\}$ from further consideration.

The problem we now confront is the dual minimization of

$$\text{MSE}[s(t), w(t)] = \sigma_a^2 \left[\sum_{n=-\infty}^{0} r_n^2 - 2r_0 + 1 + \frac{\sigma^2}{\sigma_a^2} \right]$$

with respect to $s(t)$ and $w(t)$ when a constraint is imposed on the average transmitted power.

The above expression indicates that, under the assumption of perfect past decisions, the mean-square error is minimized by minimizing both the pulse precursors in the overall impulse response and the output noise power, while keeping r_0 close to unity.

We give a precise formulation and solution to this problem in Section III.

III. RECEIVER OPTIMIZATION

Writing the MSE in detail we obtain

$$\frac{\text{MSE}}{\sigma_a^2} = 1 + \sum_{n=-\infty}^{0} \left[\int_{-\infty}^{\infty} w(\tau)p(nT - \tau)d\tau \right]^2 - 2 \int_{-\infty}^{\infty} w(\tau)p(-\tau)d\tau + \frac{N_0}{2\sigma_a^2} \int_{-\infty}^{\infty} w^2(\tau)d\tau, \quad (1)$$

where

$$p(t) = s(t) * h(t).$$

Keeping $s(t)$ fixed, and using a standard calculus-of-variation approach, results in an integral equation for $w(t)$

$$p(-t) = N_0' w(t) + \int_{-\infty}^{\infty} w(\tau) \left(\sum_{n=-\infty}^{0} p(nT - \tau)p(nT - t) \right) d\tau, \quad (2)$$

where

$$N_0' = \frac{N_0}{2\sigma_a^2}.$$

If in eq. (2) we set

$$U_n = \int_{-\infty}^{\infty} w(\tau)p(nT - \tau)d\tau,$$

we see that the optimum solution must have a representation in the form

$$w(t) = \sum_{n=-\infty}^{0} g_n p(nT - t), \tag{3}$$

where

$$g_0 = \frac{1}{N_0'}(1 - U_0)$$

and

$$g_n = -\frac{U_n}{N_0'}, \qquad n = \leqq -1.$$

In (3) is revealed the structure of the optimum receiving filter. It is composed of a *matched filter having impulse response* $p(-t)$ *followed by a one-sided (anticausal) tapped delay line with weights equal to* g_n.

Linear equations involving the set $\{U_n\}$ can be obtained by first multiplying both sides of (2) by $p(kT - t)$, $k \leqq 0$, then integrating from minus to plus infinity. The resulting linear system of equations is

$$R_k = N_0'U_k + \sum_{n=-\infty}^{0} R_{n-k}U_n, \qquad k = 0, -1\cdots, \tag{4}$$

where

$$R_k = R_{-k} = \int_{-\infty}^{\infty} p(-t)p(kT - t)dt.$$

The system of eqs. (4) can be solved by standard Wiener-Hopf techniques and the details are given in Appendix A. The solution in terms of the discrete Fourier transform of the sequence $\{U_n\}_{-\infty}^{0}$ is

$$U(\theta) = \sum_{n=-\infty}^{0} U_n e^{in\theta}$$

$$= 1 - \frac{N_0'}{M^-(\theta)\gamma_0}, \tag{5}$$

where

$$M(\theta) = M^{+}(\theta)M^{-}(\theta) = R(\theta) + N_0' = \sum_{n=-\infty}^{n=\infty} M_n e^{in\theta},$$

$$M^{+}(\theta) = \sum_{n=0}^{\infty} \gamma_n e^{in\theta},$$

and

$$M^{-}(\theta) = M^{+}(-\theta).$$

This is standard procedure making use of the well-known factorization property of covariance functions. Methods for obtaining the sequence $\{\gamma_n\}_0^\infty$ from the given sequence $\{R_n\}_{-\infty}^{\infty}$ are well documented in the literature. One method is summarized in Appendix A.

Having specified the optimum receiving filter, we now obtain a formula for the minimized mean-square error. The availability of this simple formula will allow further optimization of the transmitting filter.

Let $w_0(t)$ be the impulse response of the optimum receiving filter. [This function solves the integral equation (2).] Substitute $w_0(t)$ into (2), multiply both sides by $w_0(t)$, and integrate from minus infinity to plus infinity to obtain

$$\int_{-\infty}^{\infty} p(-t)w_0(t)dt = N_0' \int_{-\infty}^{\infty} w_0^2(t)dt + \sum_{n=-\infty}^{0} \left(\int_{-\infty}^{\infty} p(nT-t)w_0(t)dt\right)^2. \quad (6)$$

Putting this into (1) with $w(t)$ replaced by $w_0(t)$ we get a formula for the optimized MSE

$$\text{MSE}[w_0(t)] = \sigma_a^2(1 - U_0) = N_0' \sigma_a^2 g_0. \quad (7)$$

This result was obtained by Monsen[4] but unfortunately he did not go any further. As it turns out, a much richer formula than (7) can be obtained since U_0 can be expressed directly in terms of the spectrum of the channel characteristics in cascade with the transmitting filter. To carry this further, observe from (5) that

$$\begin{aligned} U_0 &= \text{dc term in } U(\theta) \\ &= 1 - \frac{N_0'}{\gamma_0^2} \end{aligned} \quad (8)$$

and consequently

$$\text{MSE} = \sigma_a^2 \frac{N_0'}{\gamma_0^2}. \quad (9)$$

As it turns out, γ_0 is functionally related to $M(\theta)$ in a rather simple manner. This relationship can be found in the literature but the derivation is short and so we briefly outline the approacb.

Under very mild conditions on $M(\theta)$ (see Doob,[5] pp. 159–161)

$$M(\theta) = \left| \sum_{n=0}^{\infty} \gamma_n e^{in\theta} \right|^2, \tag{10}$$

where γ_0 is real and positive. Since $M(\theta) > 0$, consider

$$\int_{-\pi}^{\pi} \ln M(\theta) dt = \int_{-\pi}^{\pi} \ln \left[\gamma_0 + \sum_{n=1}^{\infty} \gamma_n e^{in\theta} \right] d\theta + \int_{-\pi}^{\pi} \ln \left[\gamma_0 + \sum_{n=1}^{\infty} \gamma_n e^{-in\theta} \right] d\theta. \tag{11}$$

When tbe ln's on tbe r.b.s. of (11) are expanded in a power series and the integrations are carried out (recognizing tbat all integrals involving powers of exp $\{in\theta\}$, $n \neq 0$, vanish) we get

$$\gamma_0^2 = \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln[R(\theta) + N_0'] d\theta \right\}, \tag{12}$$

wbere

$$R(\theta) = \sum_{n=-\infty}^{\infty} R_n e^{in\theta},$$

$$R_n = \int_{-\infty}^{\infty} |P(\omega)|^2 e^{i\omega nT} \frac{d\omega}{2\pi},$$

and

$$P(\omega) = \int_{-\infty}^{\infty} s(t) * h(t) e^{i\omega t} dt.$$

After minor algebraic manipulations and changes of variables we obtain by substituting (12) into (9)

$$\text{MSE} = \sigma_a^2 \exp \left\{ -\frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \ln[Y(\omega) + 1] d\omega \right\}, \tag{13}$$

wbere

$$Y(\omega) = \frac{1}{N_0' T} \sum_{n=-\infty}^{\infty} \left| P\left(\omega - \frac{2\pi n}{T} \right) \right|^2. \tag{14}$$

This formula, as far as can be determined, is new and its simple form will enable us in Section IV to carry out an additional optimization with respect to the transmitting filter.

It is instructive at tbis point to compare this formula with the one obtained for a linear equalizer witbout decision feedback. Berger and

Tufts[6] have found such a formula and from their paper we have that

$$(\text{MSE})_{\text{linear}} = \sigma_a^2 \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} [Y(\omega) + 1]^{-1} d\omega$$
$$= \left\langle \frac{1}{Y(\omega) + 1} \right\rangle, \tag{15}$$

where

$$\langle \cdot \rangle = \sigma_a^2 \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} [\cdot] d\omega.$$

In terms of the same notation, (13) can he put into the form

$$\text{MSE} = \exp\{-\langle \ln[Y(\omega) + 1] \rangle\} \tag{16}$$

from which we get immediately that

$$\text{MSE} \leqq \left\langle e^{-\ln[Y(\omega)+1]} \right\rangle = \left\langle \frac{1}{Y(\omega) + 1} \right\rangle = (\text{MSE})_{\text{linear}}. \tag{17}$$

As expected, the mean-square error with decision feedback is always smaller than the MSE of a linear equalizer. Comparing (15) with (16)

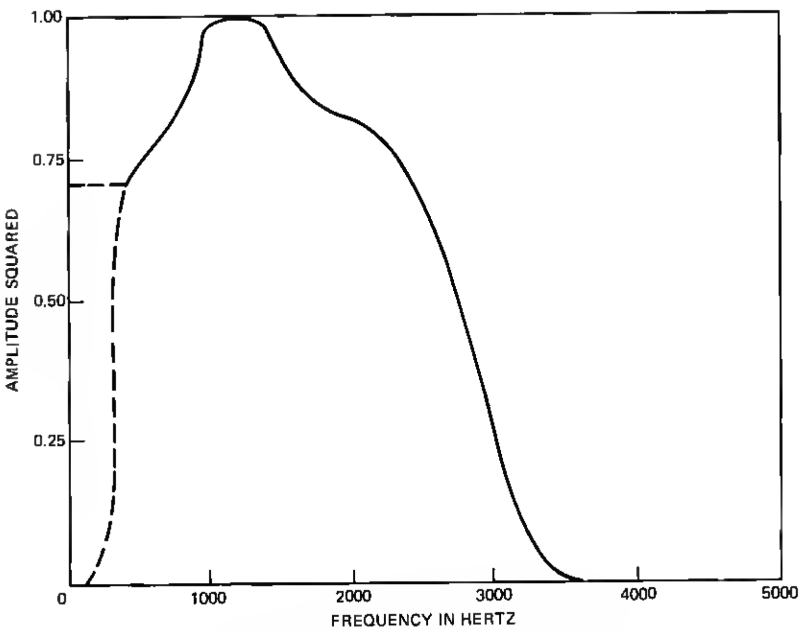

Fig. 2—Amplitude-squared characteristic of typical voiceband channel.

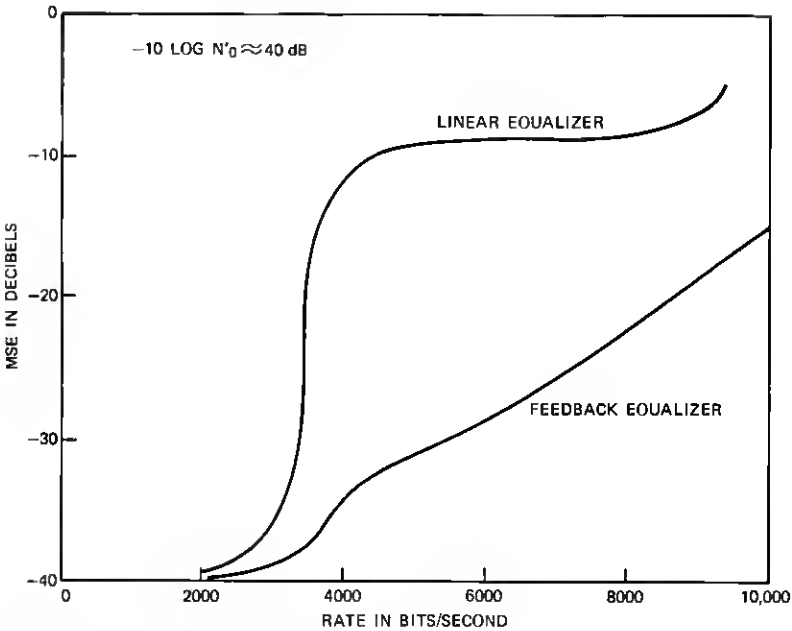

Fig. 3—MSE in dB vs binary data rate for channel shown in Fig. 2 without dc transmission.

shows that both equalization methods yield the same MSE if and only if $Y(\omega)$ is a constant, i.e., there is no intersymbol interference.

Prior to optimizing the transmitting filter we wish to illustrate the behavior of (15) and (16) for a typical voiceband channel as the signaling rate $1/T$ is allowed to increase. An amplitude-squared characteristic for a typical voiceband telephone channel is shown in Fig. 2 (the dashed line with zero transmission at zero frequency is typical). Figure 3 shows the resulting MSE vs pulse rate for both a linear equalizer and a decision feedback equalizer when $\sigma_a^2 = 1$ (binary data) and $-10 \log N_0' \approx 40$ dB. The calculations were done numerically by using (15) and (16). We note that the performance of the linear equalizer deteriorates rapidly when the rate is greater than ≈ 3000 bits/second while the decision feedback equalizer deteriorates gracefully. The reason the linear equalizer shows such a poor performance is that it must compensate for the missing energy around dc. In practice, of course, modulation is used to place the data energy at a more suitable location in the passband spectrum to avoid this severe null. To make the comparison fair, we artificially extended the channel

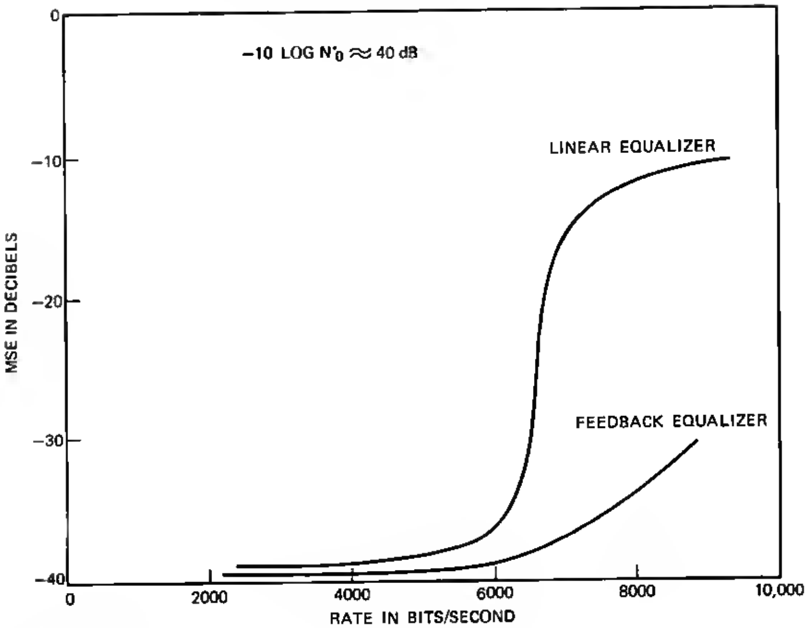

Fig. 4—MSE in dB vs binary data rate for channel shown in Fig. 2 with dc transmission.

characteristic from about 300 Hz to 0 Hz to a constant transmission. This is indicated in Fig. 2 by the dashed line parallel to the frequency axis. Figure 4 shows the comparisons for this atypical channel. As expected, the linear equalizer has a sharp threshold at approximately the Nyquist rate but, as before, the decision feedback equalizer deteriorates much more gracefully.

IV. TRANSMITTER OPTIMIZATION

The problem we address here is the optimization of (13) with respect to the transmitting filter characteristics subject to an average power constraint.

Let $S(\omega)$ and $H(\omega)$ be the Fourier transforms of $s(t)$ and $h(t)$ respectively. The average power at the output of $s(t)$ is

$$P = \frac{\sigma_a^2}{T}\int_{-\infty}^{\infty} s^2(t)dt = \frac{\sigma_a^2}{T}\frac{1}{2\pi}\int_{-\infty}^{\infty} S^2(\omega)d\omega$$
$$= \frac{\sigma_a^2}{T}\frac{1}{2\pi}\int_{-\pi/T}^{\pi/T}\left(\sum_{n=-\infty}^{\infty} S^2\left(\omega - \frac{2\pi n}{T}\right)\right)d\omega, \qquad (18)$$

where by $S^2(\omega)$ we mean $|S(\omega)|^2$.

The problem at hand is to maximize the functional

$$I = \int_{-\pi/T}^{\pi/T} \ln[K \sum_n S_n^2(\omega) H_n^2(\omega) + 1] d\omega + \lambda \int_{-\pi/T}^{\pi/T} [\sum_n S_n^2(\omega)] d\omega \quad (19)$$

with respect to the infinite set of functions $\{S_n^2(\omega), n = \text{all integers}\}$. Where

$$S_n^2(\omega) = \left| S\left(\omega - \frac{2\pi n}{T}\right)\right|^2, \qquad K = 1/N_0' T$$

and λ is a Lagrange multiplier to be determined from tbe constraint on the average transmitted power. Observe that these functions are independent over the range $-\pi/T \leqq \omega \leqq \pi/T$ and therefore consider the variation of I with respect to, say, $S_j^2(\omega)$

$$\delta_j I = \int_{-\pi/T}^{\pi/T} \left\{ \frac{KH_j^2(\omega)}{K \sum_n S_n^2(\omega) H_n^2(\omega) + 1} \delta S_j^2(\omega) + \lambda \delta S_j^2(\omega) \right\} d\omega. \quad (20)$$

When $S_j^2(\omega) \neq 0$, setting (20) to zero implies

$$\frac{KH_j^2(\omega)}{K \sum_n S_n^2(\omega) H_n^2(\omega) + 0} + \lambda = 0 \quad (21)$$

for all $\omega \in [-\pi/T, \pi/T]$. Now, assume that $H_n^2(\omega) > H_m^2(\omega)$ for $|n| < |m|$ and $\omega \in [-\pi/T, \pi/T]$. When this condition is satisfied it is not possible to solve the system of equations given in (21) unless

$$S_n^2(\omega) = 0 \qquad \text{for all } n \neq j$$

in which case we get

$$KH_j^2(\omega) = -\lambda[KS_j^2(\omega) H_j^2(\omega) + 1]. \quad (22)$$

Substituting this into (20) indicates that the largest value is obtained when $S_j^2(\omega) = S_0^2(\omega) = S^2(\omega)$ and furthermore λ must be negative.

So far we can conclude that for channels possessing monotonically decreasing amplitude characteristics, i.e., $H_m^2(\omega) > H_n^2(\omega)$, $-\pi/T \leqq \omega \leqq \pi/T$, when $|n| > |m|$, the optimum transmitting filter cuts off at the Nyquist frequency π/T. The optimum system allows no transmission outside the band $|\omega| > \pi/T$. The restrictions imposed on the cbannels are mild and are expected to be satisfied in most situations of interest. However, removing these restrictions makes the problem slightly more complicated, and it is left up to the reader to reason how it can be solved.

Making use of this partial solution permits writing the mean-square error in the simplified form

$$-\ln\left(\frac{\text{MSE}}{\sigma_a^2}\right) = \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} \ln\left[KH^2(\omega)S^2(\omega) + 1\right]d\omega \tag{23}$$

and the functional now to be further maximized with respect to the inband structure of $S^2(\omega)$ reduces to

$$I[S(\omega)] = \int_{-\pi/T}^{\pi/T} \ln\left[KS^2(\omega)H^2(\omega) + 1\right]d\omega + \lambda \int_{-\pi/T}^{\pi/T} S^2(\omega)d\omega. \tag{24}$$

As can be seen from (20) and (22), eq. (24) is maximized when

$$S^2(\omega) = \max\left[\frac{KH^2(\omega) - \lambda_0}{\lambda_1 KH^2(\omega)}, 0\right], \tag{25}$$

where $\lambda = -\lambda_0$.

To determine the Lagrange multiplier λ_0, two cases must be distinguished.

Case 1: $KH^2(\omega) - \lambda_0 > 0$, *for all* $|\omega| \leqq \pi/T$

In this case, we get

$$1 + KH^2S^2 = \mu H^2, \qquad \mu = K/\lambda_0,$$

which when substituted into (23) results in an explicit expression for the optimum MSE

$$-\ln\left(\frac{\text{MSE}}{\sigma_a^2}\right) = \ln\mu + \frac{T}{\pi}\int_0^{\pi/T} \ln H^2(\omega)d\omega. \tag{26}$$

The factor μ is determined from the average power constraint in the following manner. Use (25) and (18) to write

$$\begin{aligned} P &= \frac{\sigma_a^2}{T\pi}\int_0^{\pi/T}\left[\frac{\mu}{K} - \frac{1}{KH^2(\omega)}\right]d\omega \\ &= \frac{\sigma_a^2}{KT^2}(\mu - A), \end{aligned} \tag{27}$$

where

$$A = \frac{T}{\pi}\int_0^{\pi/T} \frac{1}{H^2(\omega)}d\omega.$$

Substituting the parameters $K = 1/TN_0'$ and $N_0' = N_0/2\sigma_a^2$ into (27) gives explicitly

$$\mu = \rho + A, \tag{28}$$

where

$$\rho = \frac{P}{\left(\frac{N_0}{2}\frac{1}{T}\right)} \equiv \frac{\text{Average transmitted signal power}}{\text{Average noise power in the Nyquist band}}.$$

Thus (26) and (28) provide a complete solution for this case.

Case 2: There exists a set of ω for which $KH^2(\omega) - \lambda_0 \leqq 0$

The optimization procedure in this case involves the standard water-pouring argument. To illustrate the nature of the solution we take the situation where $H^2(\omega)$ is strictly monotonically decreasing in the Nyquist band. This implies that there exists only one frequency ω_0 for which $KH^2(\omega) = \lambda_0$ and consequently we get

$$\mu = \frac{1}{H^2(\omega_0)}. \tag{29}$$

This gives one relation between the unknowns μ and ω_0 and another is obtained from the power constraint. Since the optimum filter characteristic is zero when $\omega > \omega_0$, the signal-to-noise ratio is

$$\rho = \frac{T}{\pi} \int_0^{\omega_0} \left[\frac{1}{H^2(\omega_0)} - \frac{1}{H^2(\omega)} \right] d\omega \tag{30}$$

and an explicit formula for the mean-square is

$$-\ln \left[\frac{\text{MSE}}{\sigma_a^2} \right] = \frac{T}{\pi} \int_0^{\omega_0} \ln H^2(\omega) d\omega - \frac{T}{\pi} \omega_0 \ln H^2(\omega_0). \tag{31}$$

We now briefly summarize how these optimized formulas are to be used:

(*i*) For a given transmitted average signal-to-noise ratio ρ, solve eq. (30) for ω_0.
(*ii*) If $\omega_0 < \pi/T$, use formula (31) to compute MSE.
(*iii*) If $\omega_0 \geqq \pi/T$, use formula (26) to compute MSE.

In Section VI we shall illustrate numerically the use of these formulas.

V. SIGNALING FASTER THAN THE NYQUIST RATE

Here we examine the behavior of the optimized mean-square error when the frequency support of an ideal unity-gain channel is smaller than the Nyquist rate $\frac{1}{2}T$. After deriving the optimized MSE for this situation, the possibility of further optimization relative to the signaling rate when the information rate per unit bandwidth is held fixed will be investigated. This has heen an open question thus far and the issue is whether increasing the signaling rate beyond the Nyquist rate while decreasing tbe number of levels to maintain a fixed information rate is ever beneficial.

Consider a channel having the characteristic

$$H^2(\omega) = \begin{cases} 0, & \omega \in E_1 \\ 1, & \omega \in E_2' \end{cases} \tag{32}$$

where the sets E_1 and E_2 form a partition on the frequency interval $E = \{\omega : 0 \leqq \omega \leqq \pi/T\}$. By frequency support we mean the measure of the set E_2 denoted hy $m(E_2)$.

For tbis channel it is easy to calculate explicitly the mean-square error. Observe from eq. (22) that for a piecewise constant channel the optimum transmitting filter is a constant wben $\omega \in E_2$ and zero otherwise. The minimum MSE is then calculated from (23)

$$\begin{aligned} -\ln\left[\frac{\text{MSE}}{\sigma_a^2}\right] &= \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} \ln\left[KS^2H^2(\omega) + 1\right]d\omega \\ &= \frac{T}{\pi}\int_{\omega \in E_2} \ln\left[KS^2 + 1\right]d\omega \\ &= \frac{T}{\pi} m(E_2)\ln\left[KS^2 + 1\right], \end{aligned} \tag{33}$$

where S^2 is a constant to be determined from the power constraint

$$P = \frac{\sigma_a^2}{\pi T}\int_{\omega \in E_2} S^2 d\omega = \frac{\sigma_a^2}{\pi T} S^2 m(E_2). \tag{34}$$

From this equation and the definition of $K = 1/(N_0'T)$ it can he checked that

$$\rho = KS^2 = \frac{\text{Average signal power}}{\text{Average noise power in a band} = m(E_2)}.$$

Suhstituting tbis into (33) gives the simple desired formula

$$\text{MSE} = \sigma_a^2(1 + \rho)^{-\alpha}, \tag{35}$$

where

$$\begin{aligned} \alpha &= \frac{m(E_2)T}{\pi} \\ &= \frac{\text{Channel bandwidtb}}{2 \times \text{Signaling rate}} \leqq 1. \end{aligned}$$

For fixed ρ, MSE $\to \sigma_a^2$ wben $\alpha = 0$ and, as expected, MSE $\to \sigma_a^2(1+\rho)^{-1}$ wben $\alpha = 1$. It is curious that as long as $\alpha \neq 0$, MSE $\to 0$ as $\rho \to \infty$.

When the set E_2 is the interval $I = \{\omega : 0 \leqq \omega \leqq \pi/T_0 < \pi/T\}$, it is referred to as being less than the Nyquist band. Since there is no mathematical reason for making this distinction we shall refer to "signaling faster than the Nyquist rate" whenever $\alpha < 1$.

We now investigate whether it is ever advantageous to signal faster than the Nyquist rate. Clearly, for fixed $\sigma_a^2 = (L^2 - 1)/3$, or a fixed number of levels, (35) shows that MSE degrades rapidly with decreasing α. An interesting question, first raised by R. W. Lucky,[7] is the possibility of trading L with α to further minimize the mean-square error. This is the problem we address.

Let the source information rate be $R = \log_2 L/T$ bits/second. The available bandwidth is equal to $1/(2\pi)m(E_2)$ cycles/second. Thus the normalized information rate is

$$\begin{aligned}\mathcal{E} &= \frac{R}{\dfrac{1}{2\pi}\, m(E_2)} = \frac{2}{\left[\dfrac{Tm(E_2)}{\pi}\right]} \log_2 L \\ &= \frac{2}{\alpha} \log_2 L \,\frac{\text{bits}}{\text{cycle}}. \end{aligned} \tag{36}$$

Writing (35) in terms of this quantity gives

$$\mathrm{MSE}(\alpha) = \frac{2^{\alpha\mathcal{E}} - 1}{3}\, \frac{1}{(1+\rho)^{\alpha}}. \tag{37}$$

Letting $C = \log_2(1 + \rho)$ be the ultimate attainable rate according to Shannon's theory, eq. (37) can be put into the form

$$\mathrm{MSE}(\alpha) = (2^{-\alpha(C-\mathcal{E})} - 2^{-\alpha C})\tfrac{1}{3}. \tag{38}$$

Note that, since $L = 2^{\alpha\mathcal{E}/2}$ and the minimum allowable L is equal to 2, the parameter α must be in the range $(2/\mathcal{E}, 1)$.

The problem initially posed can now be stated as follows. Find a set of α's, $2/\mathcal{E} \leqq \alpha \leqq 1$, which minimize eq. (38). We begin by setting the derivative of (38) to zero,

$$\frac{d\mathrm{MSE}(\alpha)}{d\alpha} = -(C - \mathcal{E})2^{-\alpha(C-\mathcal{E})} + C2^{-\alpha C} = 0,$$

from which we find a unique stationary point

$$\alpha = \frac{2}{\mathcal{E}} \log_2 \left[\frac{C}{C - \mathcal{E}}\right]^{\frac{1}{2}}. \tag{39}$$

Since $\mathrm{MSE}(0) = 0$ and $\mathrm{MSE}(\alpha) > 0$, the value of α found above must be a point where MSE attains a maximum. If this maximum is in tbe range $(0 \leqq \alpha \leqq 2/\mathcal{E})$, then the minimum value of MSE is attained at the boundary $(\alpha = 1)$. The condition for this to be the case is determined from (39).

$$\log_2 \left[\frac{C}{C - \mathcal{E}} \right]^{\frac{1}{2}} \leqq 1. \tag{40}$$

From this we deduce that as long as $\mathcal{E} < \frac{3}{4}C$, $\alpha = 1$ minimizes MSE. For this region of $\mathcal{E}$ and C there is no advantage gained by signaling faster than the Nyquist rate. Since C is channel capacity, $\mathcal{E}_c = \frac{3}{4}C$ seems to be a critical rate. If the rate is greater than this critical rate, the maximum point lies within the allowable range of $\alpha(2/\mathcal{E}, 1)$ thus raising the possibility tbat either $\alpha = 2/\mathcal{E}$ or $\alpha = 1$ is a minimum point. Suppose $\alpha = 2/\mathcal{E}$ is a minimum point. Equation (38) gives for this case

$$\text{MSE}\left(\alpha = \frac{2}{\mathcal{E}}\right) = 2^{-2C/\mathcal{E}} \geqq 2^{-8/3} \approx 0.157. \tag{41}$$

Tbus we have found a region for which signaling faster than the Nyquist rate appears to be beneficial. But at this high level of MSE, we are no longer justified in assuming tbat the feedback decisions are correct most of the time. In fact, what is more likely to happen is that errors begin to occur resulting in a larger value of MSE than predicted by the error-free model. We are therefore led to the conclusion that a minimum point other than at $\alpha = 1$ will render an MSE to be outside the range of practical utility. To emphasize this point further substitute (41) into (38) to get

$$\text{MSE}(\alpha = 1) = [\text{MSE}(2/\mathcal{E})]^{\mathcal{E}/2} \frac{2^{\mathcal{E}} - 1}{3}. \tag{42}$$

Tbus $\alpha = 1/\mathcal{E}$ is a minimum point whenever

$$[\text{MSE}(2/\mathcal{E})]^{\mathcal{E}/2} \left(\frac{2^{\mathcal{E}} - 1}{3}\right) > \text{MSE}(2/\mathcal{E})$$

or

$$\text{MSE}(2/\mathcal{E}) > \left(\frac{3}{2^{\mathcal{E}} - 1}\right)^{2/\mathcal{E} - 2}.$$

As an example of the use of tbese inequalities, suppose tbat $\mathcal{E} = 3$ and $C = 3.5$; therefore, $\alpha = \frac{2}{3}$ is the minimum point and tbe achievable MSE $= 0.198$ which can be obtained with a binary system $(L = 2)$. On tbe other hand, $M(\alpha = 1) = 2^{-3.5}(7/3) \cong 0.206$ which can be achieved witb $(L = 2^{1.5} = 2.8)$!!

It is interesting to see what MSE can be achieved when only a linear equalizer is used. Using formula (15) and following the same reasoning as before yields an expression for the optimized mean-square error

$$(\text{MSE})_{\text{linear}} = \sigma_a^2 \frac{(1-\alpha)\rho + 1}{\rho + 1}. \tag{43}$$

In this case, unlike in the decision feedback case [eq. (35)], as $\rho \to \infty$, MSE $\to 1 - \alpha$, when $\alpha > 0$. Thus the mean-square error cannot be made vanishingly small as the signal-to-noise ratio increases without bound.

Expressing (43) in terms of the normalized rate $\mathcal{E}$ we get

$$(\text{MSE})_{\text{linear}} = \frac{2^{\alpha\mathcal{E}} - 1}{3}\left[\frac{\rho(1-\alpha)+1}{\rho+1}\right]. \tag{44}$$

Again we seek to minimize (44) with respect to α in the range $(2/\mathcal{E}, 1)$. It can be checked that (44) has at most one stationary point in the range $0 \leqq \alpha \leqq 1$. Since $\alpha = 0$ is a minimum point, the minimum in the range $(2/\mathcal{E}, 1)$ must lie on the boundary. Thus the condition for achieving a smaller MSE when signaling faster than the Nyquist rate $(\alpha < 1)$ is

$$\rho(1 - 2/\mathcal{E}) + 1 \leqq \frac{2^{\mathcal{E}} - 1}{3}$$

or

$$\rho = 2^C - 1 \leqq \frac{2^{\mathcal{E}} - 1}{3} \frac{1}{1 - 2/\mathcal{E}}. \tag{45}$$

Is it possible to find an $\mathcal{E} \leqq C$, $\mathcal{E} > 2$ which satisfies the inequality (45)? A straightforward analysis reveals that the answer is negative. In other words, the Nyquist rate is optimum provided the information rate is less than channel capacity.

VI. APPLICATION TO A CABLE CHANNEL†

This section will illustrate the use of the formulas developed in previous sections in a particular application. For this purpose we choose a cable channel having frequency characteristic

$$H(f) = \exp\{\sqrt{-2i\alpha f}\}. \tag{46}$$

We shall develop in detail the applicable formulas, provide asymptotic behaviors, and exhibit numerically the relevant parameter trade-offs. For comparison purposes, the applicable formulas for the optimum linear equalizer will also be developed.

† I am indebted to Dr. Robert Price for calling my attention to Ref. 8 where related work is reported. The paper is in Japanese; however, Dr. Price has an English translation.

We begin by first considering a suboptimum system where the transmitting filter is flat across the Nyquist band and zero outside. For this case, the minimum attainable mean-square error is [eq. (23)]

$$M_d = \exp\left\{-\frac{T}{\pi}\int_0^{\pi/T} \ln\left[S^2Ke^{-\sqrt{2\omega\alpha/\pi}} + 1\right]d\omega\right\}, \tag{47}$$

where

$$M_d = \mathrm{MSE}/\sigma_a^2, \qquad |H(\omega)|^2 = e^{-\sqrt{2\omega\alpha/\pi}}, \qquad \text{and} \quad K = 2\sigma_a^2/TN_0$$

are to be determined from the average power constraint. Since the transmitted power $P = (\sigma_a^2/T)S^2$, we find that $S^2K = 2PT/N_0 = \rho$, the transmitted signal-to-noise ratio, where the noise is measured in a band $= 1/T$. Substituting these constants into (47) and making some changes of variables result in

$$M_d = \exp\left\{-2\int_0^{\frac{1}{2}} \ln\left[\rho e^{-\sqrt{4\beta y}} + 1\right]dy\right\}, \tag{48}$$

where $\beta = \alpha/T$. The parameter $\sqrt{2\beta}$ is seen to be proportional to the loss of the cable in dB at the Nyquist frequency $\frac{1}{2}T$.

We are interested in the behavior of (48) when ρ and β are varied. While it is not possible to express this integral in a closed form, it is possible to obtain a rapidly converging series in the two parameters of interest from which asymptotic behaviors can be deduced. Appendix B shows the details of the development. Different power series apply in different regions. The first series applies when $\ln\rho < \sqrt{2\beta}$ and the second $\ln\rho \geqq \sqrt{2\beta}$. The results are as follows

1: $\sqrt{2\beta} \geqq \ln\rho > 0$

$$-\ln M_d = \frac{(\ln\rho)^3 + \pi^2\ln\rho}{6\beta} - \sqrt{\frac{2}{\beta}}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{[\rho e^{-\sqrt{2\beta}}]^n}{n^2} + \frac{1}{\beta}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^3}\left[\frac{1}{\rho^n} - (\rho e^{-\sqrt{2\beta}})^n\right]. \tag{49}$$

2: $\ln\rho \geqq \sqrt{2\beta} > 0$

$$-\ln M_d = \ln\rho - \frac{2}{3}\sqrt{2\beta} + \sqrt{\frac{1}{\beta}}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^2}\left[\frac{e^{\sqrt{2\beta}}}{\rho}\right]^n + \frac{1}{\beta}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^3}\left[\frac{1}{\rho^n} - \left(\frac{e^{\sqrt{2\beta}}}{\rho}\right)^n\right]. \tag{50}$$

It can be checked that (49) equals (50) when $\ln\rho = \sqrt{2\beta}$. The first asymptotic behavior is deduced from (50) when $\rho \to \infty$ and $\sqrt{2\beta}$ is held fixed. For this case we get

$$M_d \sim \frac{e^{\frac{1}{2}\sqrt{2\beta}}}{\rho}. \tag{51}$$

Another asymptotic behavior is deduced from (49) as $\beta \to \infty$ while ρ is held fixed. In this case

$$M_d \sim e^{-g(\rho)/\beta}, \tag{52}$$

where

$$g(\rho) = \frac{(\ln \rho)^3 + \pi^2 \ln \rho}{6} + \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^3} \frac{1}{\rho^n}.$$

In order to make possible performance comparisons, we develop similar formulas and asymptotes for a system employing only a linear equalizer. The minimum mean-square error applicable in this situation is obtained from eq. (15). After substituting the cable characteristic we obtain

$$M_L = 2 \int_0^{\frac{1}{2}} [\rho e^{-\sqrt{4\beta y}} + 1]^{-1} dy, \tag{53}$$

where

$$M_L = \frac{(\mathrm{MSE})_{\mathrm{linear}}}{\sigma_a^2}.$$

Here, as in the decision feedback case, rapidly converging series can be developed from which asymptotic formulas are deduced. The detailed calculations are also given in Appendix B. The desired results are

1: $\sqrt{2\beta} \geqq \ln \rho > 0$

$$M_L = 1 - \frac{3(\ln \rho)^2 + \pi^2}{6\beta} + \sqrt{\frac{2}{\beta}} \ln [1 + \rho e^{-\sqrt{2\beta}}] + \frac{1}{\beta} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2} \left[\frac{1}{\rho^n} + (\rho e^{-\sqrt{2\beta}})^n \right]. \tag{54}$$

2: $\ln \rho > \sqrt{2\beta} > 0$

$$M_L = \sqrt{\frac{2}{\beta}} \ln \left[1 + \frac{e^{\sqrt{2\beta}}}{\rho} \right] + \frac{1}{\beta} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2} \left[\frac{1}{\rho^n} - \left(\frac{e^{\sqrt{2\beta}}}{\rho} \right)^n \right]. \tag{55}$$

The asymptotic formulas are readily deduced from (54) and (55).

When $\rho \to \infty$ and β is fixed we get

$$M_L \sim Q(\beta)\frac{1}{\rho}, \tag{56}$$

where

$$Q(\beta) = e^{\sqrt{2\beta}}\left[\sqrt{\frac{2}{\beta}} - \frac{1}{\beta} + \frac{e^{-\sqrt{2\beta}}}{\beta}\right].$$

On the other hand, when $\beta \to \infty$ and ρ is kept fixed,

$$M_L \sim 1 - \frac{3(\ln \rho)^2 + \pi^2}{6\beta} + \frac{D}{\beta}, \tag{57}$$

where

$$D = \sum_{n=1}^{\infty} (-1)^{n+1}\frac{1}{n^2}\frac{1}{\rho^n}.$$

At this point it is possible to make some judicious performance comparisons between the two schemes. The first observation is that, in order to get a small mean-square error, ρ must be large and greater than $e^{\sqrt{2\beta}}$. In tbis case, (51) and (56) apply. On the other hand, when $\sqrt{2\beta} > \ln \rho$, and $\beta \to \infty$, performance deteriorates rapidly as can be seen from (52) and (57). Suppose now that a large signal-to-noise ratio is available and we wish to obtain the same mean-square error in both schemes. How do the signaling rates compare?

Equating (51) and (56) shows that $\beta_d/\beta_L \sim 9/4$ when these quantities are large. In other words, asymptotically, the signaling speed of the cable may be increased by more than a factor of two with the use of decision feedback. Clearly when β is small no significant advantage can be obtained from using decision feedback equalization.

To exhibit these phenomena further, we have used numerical integration to evaluate (49) and (53) and checked the accuracy by summing terms in the various power series. The results of these calculations are exhibited graphically in Figs. 5 through 9. A striking feature in all these curves is the manner MSE degrades as β increases. The linear MSE exhibits a sharp threshold while the MSE for the decision feedback equalizer degrades much more gracefully.

Next we wish to examine the possible payoffs when the inband characteristics of the transmitter filter are optimized. To do this explicitly, we follow the procedure outlined in Section IV. Equation (30) must first be evaluated for the cable characteristics. (We omit all straightforward integrations and algebraic manipulations.) Re-

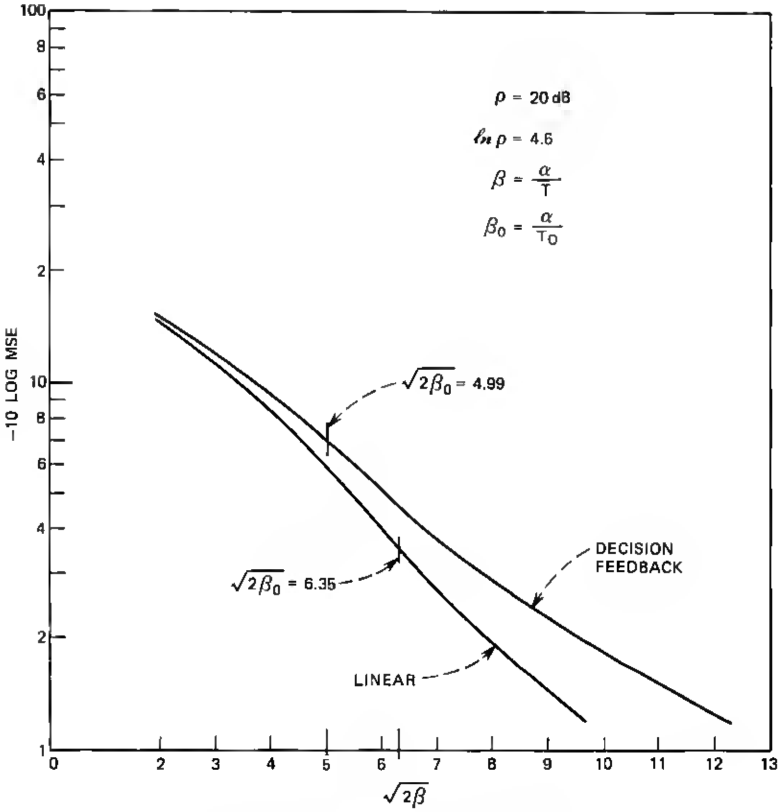

Fig. 5—MSE in dB vs $\sqrt{2\beta}$ for $\rho = 20$ dB.

writing eq. (30),

$$\rho = \frac{P}{\left(\frac{N_0}{2}\frac{1}{T}\right)} = \frac{T}{\pi}\int_0^{\omega_0}\left[\frac{1}{H^2(\omega_0)} - \frac{1}{H^2(\omega)}\right] d\omega$$

$$= \frac{\beta_0}{\beta}\left[e^{\sqrt{2\beta_0}} - \frac{F(2\beta_0)}{\beta_0}\right], \tag{58}$$

where

$$\beta_0 = \alpha/T_0, \qquad \omega_0 = \pi/T_0, \qquad \beta = \alpha/T,$$

$$H^2(\omega) = e^{-\sqrt{2\omega\alpha/\pi}},$$

and

$$F(x) = e^{\sqrt{x}}(\sqrt{x} - 1) + 1 \sim x/2 \text{ when } x \text{ is small.}$$

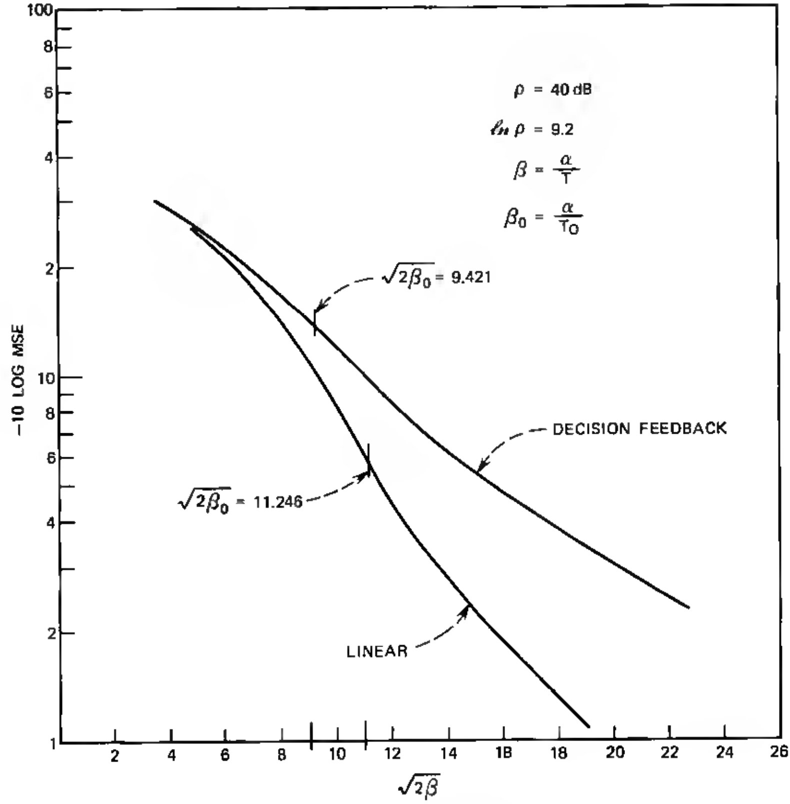

Fig. 6—MSE in dB vs $\sqrt{2\beta}$ for $\rho = 40$ dB.

The explicit evaluation of MSE is as follows: For a given ρ, β solve (58) for β_0. If $\beta_0 > \beta$, calculate MSE from eq. (26),

$$-\ln\left[\frac{\text{MSE}}{\sigma_a^2}\right] = \ln M_d = \ln \mu + \frac{T}{\pi}\int_0^{\pi/T} \ln H^2(\omega)d\omega.$$

An explicit evaluation gives

$$M_d = \frac{e^{\frac{1}{2}\sqrt{2\beta}}}{\rho + \dfrac{F(2\beta)}{\beta}}, \qquad \beta_0 \geq \beta. \tag{59}$$

On the other hand, if (58) yields a $\beta_0 < \beta$, use formula (31) to compute

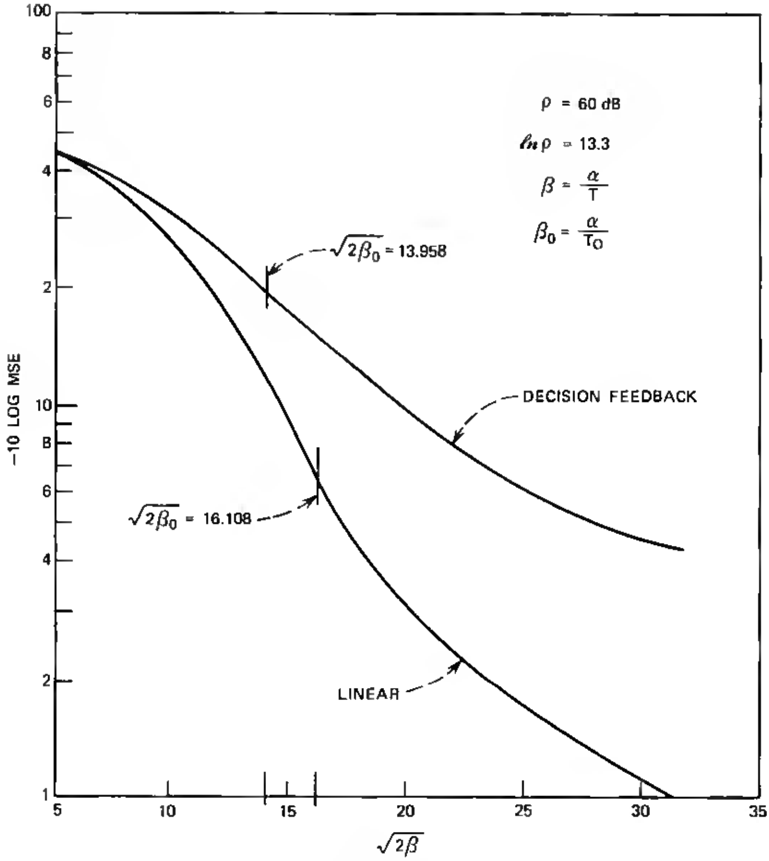

Fig. 7—MSE in dB vs $\sqrt{2\beta}$ for $\rho = 60$ dB.

MSE. The explicit evaluation for this case gives

$$M_d = e^{-\frac{1}{2}(\beta_0/\beta)\sqrt{2\beta_0}}, \qquad \beta_0 \geq \beta, \tag{60}$$

$$= \left[\frac{1}{\rho \dfrac{\beta}{\beta_0} + F(2\beta_0)/\beta_0} \right]^{\frac{1}{2}(\beta_0/\beta)},$$

where $e^{\sqrt{2\beta_0}}$ was obtained from (58). It can be checked that when $\beta_0 = \beta$ in (58), eq. (59) equals (60) as it must.

Let us now pause and examine what these optimized results are telling us. Suppose β is fixed in (58) and ρ is allowed to increase.

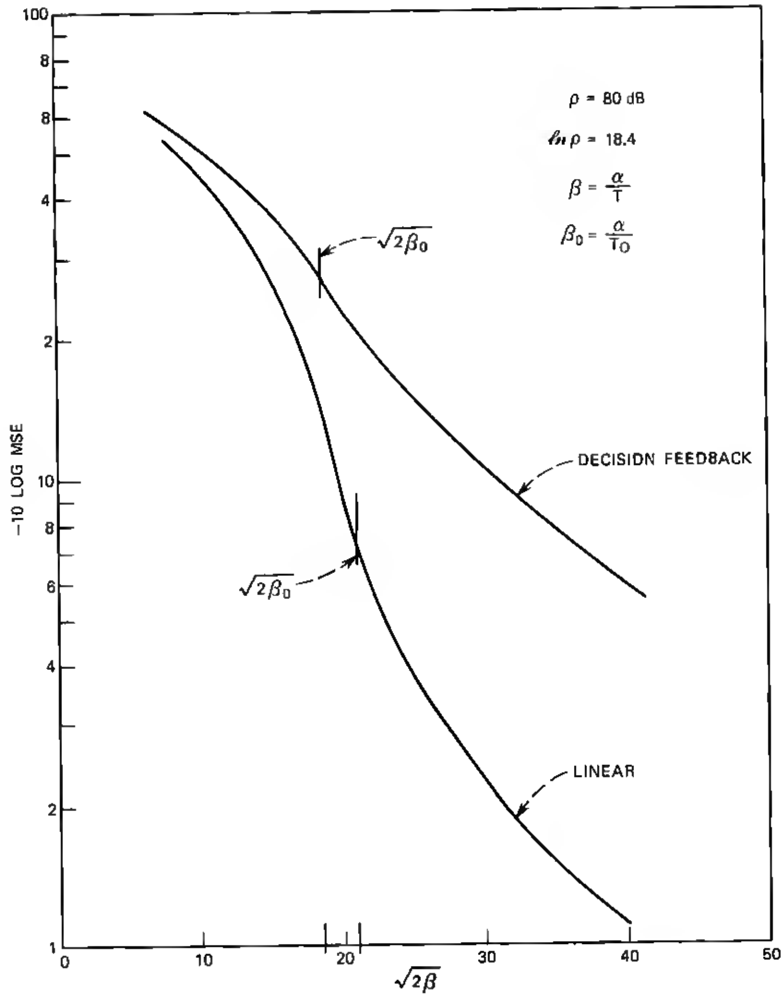

Fig. 8—MSE in dB vs $\sqrt{2\beta}$ for $\rho = 80$ dB.

Eventually a β_0 will be found which satisfies (58) and which ultimately will be greater than β. The physical implication of finding a β_0 which is less than β is that the transmitting filter cuts off before the Nyquist frequency $\frac{1}{2}T$. This will occur only when ρ is relatively small and thus results in a poor MSE. Practically, the region of interest is when ρ is large such that $\beta_0 \geqq \beta$, in which case the filter cuts off at the Nyquist frequency. In this region (59) applies and, upon comparing (59) with

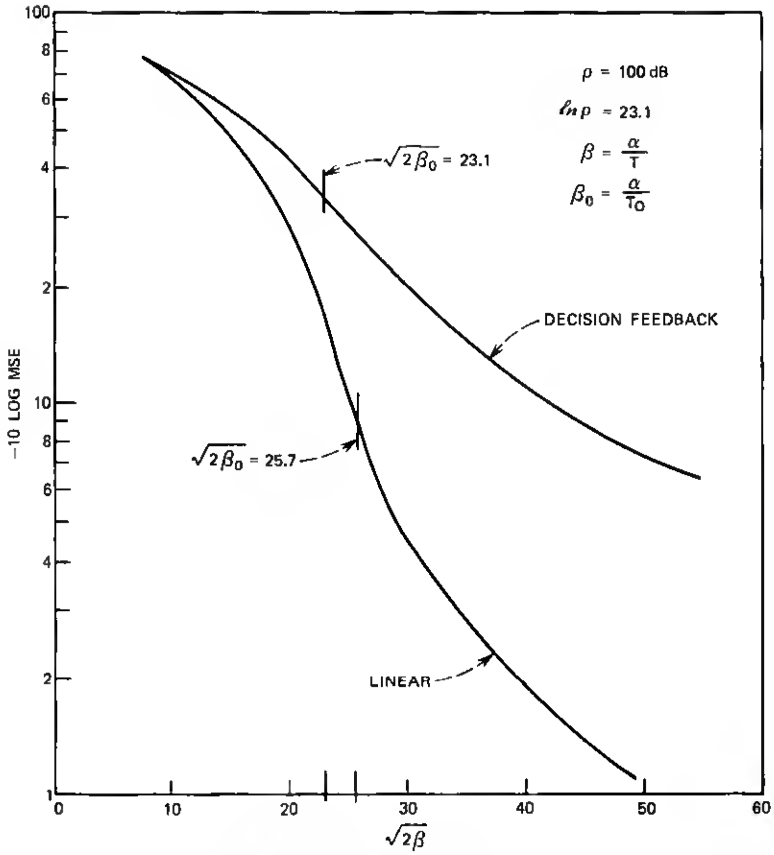

Fig. 9—MSE in dB vs $\sqrt{2\beta}$ for ρ = 100 dB.

the asymptotic suboptimized result, (51) shows that

$$\frac{e^{\frac{1}{2}\sqrt{2\beta}}}{\rho + \dfrac{F(2\beta)}{\beta}} \leqq \frac{e^{\frac{1}{2}\sqrt{2\beta}}}{\rho} .$$

Since $\min_\beta [F(\beta)/\beta] = 1$, the optimized result appears to he asymptotically equal to the suboptimized result. In other words, in the region where $\ln \rho > \sqrt{2\beta}$ and $\rho \rightarrow \infty$ no henefits are ohtained from inhand optimization. This is also evident from eq. (25) since when $\ln \rho$ is large relative to $\sqrt{2\beta}$ the structure of the optimum transmitting filter is a constant. The situation where $\beta_0 < \beta$ is slightly more complicated to compare. Here inband optimization should perhaps be beneficial. However, comparisons in this case between the optimized MSE and the suboptimized ones must be made on the basis of the same transmitted power rather than signal-to-noise ratio because the systems operate over different bandwidths.

Again for comparison purposes, we summarize the formulas that apply when the inband characteristics of the transmitting filter in a system using only linear equalization are optimized. Berger and Tufts[6] carried out such an optimization and the procedure is similar to the one carried out in Section V. Adopting our notation and the same definition of parameters as above we can obtain explicitly the following formulas applicable for a linear system.

Choose a ρ and a β and solve for β_0 in the equation below:

$$\rho = \frac{\beta_0}{\beta}\left[\frac{4}{\beta_0}F(\beta_0/2)e^{\sqrt{\beta_{02}}} - \frac{1}{\beta_0}F(2\beta_0)\right]. \tag{61}$$

If $\beta_0 > \beta$, calculate

$$M_L = \frac{\left[\dfrac{4}{\beta}F(\beta/2)\right]^2}{\rho + \dfrac{F(2\beta)}{\beta}}, \qquad \beta_0 > \beta. \tag{62}$$

If β_0 in (61) is $< \beta$, calculate

$$M_L = 1 - \frac{\beta_0}{\beta} + \frac{4}{\beta}F(\beta_0/2)e^{-\sqrt{\beta_0/2}}. \tag{63}$$

It is now possible to cross plot the formulas derived in this section *ad nauseum*. We shall show only two sets of graphs. Figures 10 and 11 show four curves of MSE/σ_a^2 in dB vs $\sqrt{2\beta}$ where $E = 2\alpha P/N_0$ and P is the transmitted power divided by the parameter $N_0/2\alpha$. In each case we plot the optimized results and the suboptimized results. The optimized decision feedback equalizer results were evaluated from equations (58), (59), and (60) and from equations (61), (62), and (63) for the linear equalizer. The nonoptimized results are given in (48) and (53) respectively. In all cases $E = \rho/\beta$ in dB. Marked on the curves is the value of $\sqrt{2\beta_0}$ where the transmitting filters cut off. We show two cases, $10 \log_{10}E = 60$ and $10 \log_{10}E = 100$. It appears that inband optimization does not provide a great deal of performance enhancement. As expected, inband optimization yields more improvement in the linear equalization scheme than in decision feedback.

We have also evaluated the optimized results as a function of the actual signal-to-noise ratio ρ, where the noise is measured in whatever

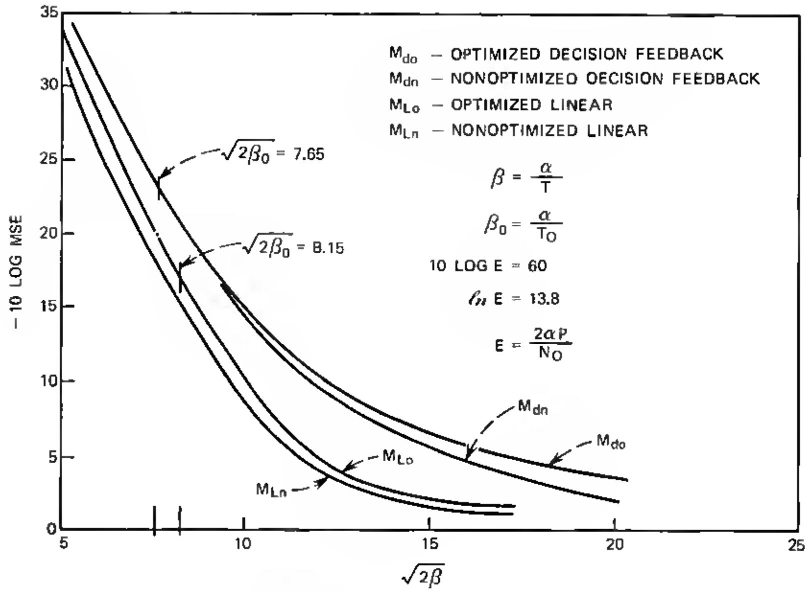

Fig. 10—MSE in dB vs $\sqrt{2\beta}$ for $E = 60$ dB.

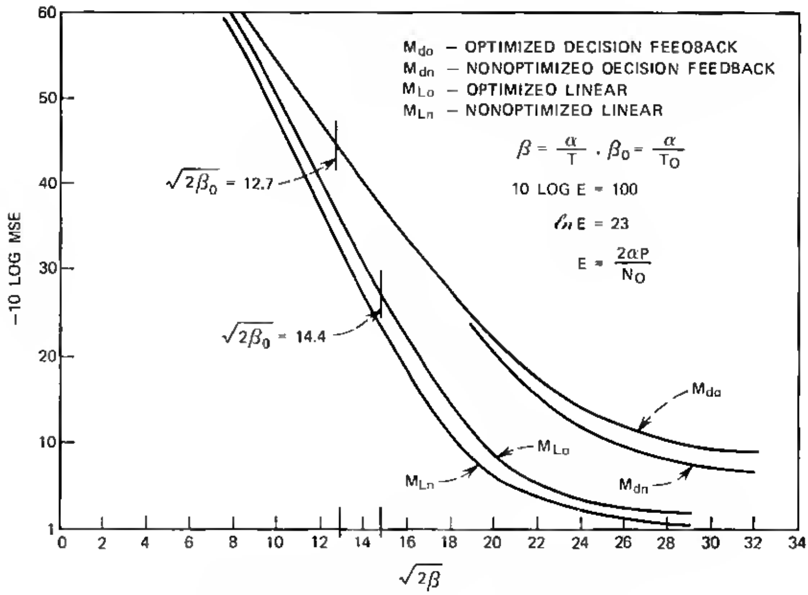

Fig. 11—MSE in dB vs $\sqrt{2\beta}$ for $E = 100$ dB.

band happens to be optimum. We found insignificant differences between these and the suboptimized results shown in Figs. 5 through 10.

In concluding this section we wish to stress that in practice error propagation problems may negate the indicated theoretical results for this channel. When the MSE is large, errors will result. In addition, the tap gains of the feedback filter may become quite large causing those errors which do result to propagate.

VII. ACKNOWLEDGMENTS

During the course of this work I have spoken to many people about this problem and received valuable suggestions. I wish to name these people in alphabetical order: L. H. Brandenburg, D. D. Falconer, G. J. Foschini, R. D. Gitlin, S. Halfin, R. W. Lucky, F. R. Magee, J. E. Mazo, S. O. Rice, and M. Segal. Finally, a great deal of credit is due to R. R. Anderson who did all the numerical work in this paper.

APPENDIX A

Solution of the Wiener-Hopf Equations

We wish to solve the set of linear equations

$$R_k = \sum_{n=-\infty}^{0} M_{n-k} S_n, \qquad k = 0, -1, -2 \cdots, \tag{64}$$

where $\{R_k\}_{-\infty}^{\infty}$ and $M_{n-k} = R_{n-k} + N_0' \delta_{n-k}$ ($\delta_{n-k} = 1, n = k; \delta_{n-k} = 0, n \neq k$) are given.

Since $\{M_n\}_{-\infty}^{\infty}$ is a correlation sequence with positive Fourier coefficients it is well known that it can be represented as the discrete convolution of a sequence $\{M_n^-\}_{-\infty}^{0}$ and a sequence $\{M_n^+\}_0^{\infty}$, namely

$$M_n = \sum_{j=0}^{\infty} M_j^+ M_{n-j}^- \qquad \text{for all } n. \tag{65}$$

Let the sequence $\{X_n\}_{-\infty}^{\infty}$ be determined from

$$R_k = \sum_{j=0}^{\infty} M_j^+ X_{k-j}, \qquad \text{all } k. \tag{66}$$

Substituting (65) and (66) into (64) gives

$$\sum_{j=0}^{\infty} M_j^+ \left\{ X_{k-j} - \sum_{n=-\infty}^{0} S_n M_{n-k-j}^- \right\} = 0. \tag{67}$$

Clearly a solution of

$$X_k = \sum_{n=-\infty}^{0} S_n M_{n-k}^{-}, \qquad k \leqq 0, \tag{68}$$

is also a solution of (67).

Define the two-sided discrete Fourier transform of a sequence $\{X_n\}_{-\infty}^{\infty}$ by

$$X(\theta) = \sum_{n=-\infty}^{\infty} X_n e^{in\theta}.$$

Take the transform of both sides of (66) to obtain

$$R(\theta) = M^{+}(\theta)X(\theta). \tag{69}$$

The one-sided transform $\left(\sum_{n=-\infty}^{0} \right)$ of (68) is

$$X^{-}(\theta) = S^{-}(\theta)M^{-}(\theta) \tag{70}$$

and $X^{-}(\theta)$ is obtained from (69) as

$$X^{-}(\theta) = \left[\frac{R(\theta)}{M^{+}(\theta)} \right]_{-}, \tag{71}$$

where $[\cdot]_{-}$ stands for "projection to negative integers only." To obtain the projection, expand $[\cdot]$ in a two-sided Fourier series and retain only the part of the series containing negative/positive coefficients (including zero).

Thus the desired solution is

$$S^{-}(\theta) = \frac{1}{M^{-}(\theta)} \left[\frac{R(\theta)}{M^{+}(\theta)} \right]_{-}. \tag{72}$$

To proceed further, observe that since $M(\theta) = M^{+}(\theta)M^{-}(\theta)$ and $M(\theta) = R(\theta) + N_0'$ it is possible to calculate explicitly

$$\left[\frac{R(\theta)}{M^{+}(\theta)} \right]_{-} = M^{-}(\theta) + \frac{N_0'}{\gamma_0}, \tag{73}$$

where γ_0 is the dc coefficient of $M^{+}(\theta)$.

The final solution for $S^{-}(\theta)$ is therefore

$$S^{-}(\theta) = 1 - \frac{N_0'}{M^{-}(\theta)\gamma_0}. \tag{74}$$

As pointed out in the text, there are various methods available for calculating $M^{\pm}(\theta)$ from a known function $M(\theta)$. We briefly outline one such approach. Since $M(\theta) > 0$, $0 \leqq \theta \leqq 2\pi$, $\ln M(\theta)$ may be expanded in a two-sided Fourier series

$$\ln M(\theta) = \sum_{n=-\infty}^{0} \gamma_n^- e^{in\theta} + \sum_{n=0}^{\infty} \gamma_n^+ e^{in\theta}. \tag{75}$$

Knowing the sequence $\{\gamma_n^\pm\}_{-\infty}^{\infty}$ we can get immediately

$$M^+(\theta) = \exp\left\{\sum_{n=0}^{\infty} \gamma_n^+ e^{in\theta}\right\} \tag{76}$$

and

$$M^-(\theta) = \exp\left\{\sum_{n=-\infty}^{0} \gamma_n^- e^{in\theta}\right\}.$$

Notice that the dc term of $M^+(\theta)$ equals the dc term of $M^-(\theta)$.

APPENDIX B

Evaluation of Integrals

B.1 *Decision Feedback*

The detailed evaluation of

$$I = 2\int_0^{\frac{1}{2}} \ln\left[1 + \rho e^{-\sqrt{4\beta y}}\right] dy \tag{77}$$

is accomplished as follows: Change the variable of integration to $x = (\sqrt{4\beta y} - \ln \rho)$ which gives

$$I = \frac{1}{\beta}\int_{-\ln \rho}^{\sqrt{2\beta}-\ln \rho} \ln[1 + e^{-x}][x + \ln \rho] dx. \tag{78}$$

Assume $\sqrt{2\beta} > \ln \rho > 0$ and write (78) as

$$\begin{aligned} I &= \frac{1}{\beta}\left(\int_{-\ln \rho}^{0} + \int_0^{\sqrt{2\beta}-\ln \rho}\right) \\ &= \frac{1}{\beta}\int_0^{\ln \rho} x(\ln \rho - x) dx + \frac{1}{\beta}\int_0^{\ln \rho} (\ln \rho - x) \ln[1 + e^{-x}] dx \\ &\qquad + \frac{1}{\beta}\int_0^{\sqrt{2\beta}-\ln \rho} (x + \ln \rho) \ln[1 + e^{-x}] dx \qquad (79) \\ &= \frac{(\ln \rho)^3}{2\beta} - \frac{(\ln \rho)^3}{3\beta} + \frac{1}{\beta}\int_0^{\ln \rho} (\ln \rho - x) \ln[1 + e^{-x}] dx \\ &\qquad + \frac{1}{\beta}\int_0^{\sqrt{2\beta}-\ln \rho} (x + \ln \rho) \ln(1 + e^{-x}) dx. \end{aligned}$$

Since $x > 0$ in the range of integration, expand

$$\ln [1 + e^{-x}] = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{e^{-nx}}{n}$$

and substitute into (79) to obtain

$$I = \frac{(\ln \rho)^3}{6\beta} + \frac{1}{\beta} \sum_{n=1}^{\infty} (-1)^{n+1} \left\{ \frac{\ln \rho}{n} A_n(\ln \rho) - \frac{B_n(\ln \rho)}{n} + \frac{B_n(\sqrt{2\beta} - \ln \rho)}{n} + \frac{\ln \rho}{\beta} \frac{A_n(\sqrt{2\beta} - \ln \rho)}{n} \right\}, \quad (80)$$

where

$$A_n(\xi) = \frac{1}{n}(1 - e^{-n\xi})$$

and

$$B_n(\xi) = \frac{1 - e^{-n\xi} - n\xi e^{-n\xi}}{n^2}.$$

Collecting terms and recognizing that

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2} = \frac{\pi^2}{12}$$

we finally get

$$I = \frac{(\ln \rho)^3 + \pi^2 \ln \rho}{6\beta} - \sqrt{\frac{2}{\beta}} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(\rho e^{-\sqrt{2\beta}})^n}{n^2} + \frac{1}{\beta} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^3} \left[\frac{1}{\rho^n} - (\rho e^{-\sqrt{2\beta}})^n \right]. \quad (81)$$

When $\ln \rho > \sqrt{2\beta} > 0$, (77) can be expressed in the form

$$\begin{aligned} I &= \frac{1}{\beta} \int_{-\ln \rho}^{0} (x + \ln \rho) \ln [1 + e^{-x}] dx + \frac{1}{\beta} \int_{0}^{-(\ln \rho - \sqrt{2\beta})} (x + \ln \rho) \ln [1 + e^{-x}] dx \\ &= \frac{1}{\beta} \int_{0}^{\ln \rho} (\ln \rho - x)[x + \ln(1 + e^{-x})] dx - \frac{1}{\beta} \int_{0}^{\ln \rho - \sqrt{2\beta}} (\ln \rho - x)[x + \ln (1 + e^{-x})] dx. \quad (82) \end{aligned}$$

At this stage $\ln (1 + e^{-x})$ is again expanded in a power series and when the terms are collected we obtain

$$I = \ln \rho - \frac{2}{3} \sqrt{2\beta} + \sqrt{\frac{2}{\beta}} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{e^{n\sqrt{2\beta}}}{n^2 \rho^n} + \frac{1}{\beta} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^3 \rho^n} [1 - e^{n\sqrt{2\beta}}]. \quad (83)$$

B.2 *Linear Equation*

The integral we wish to evaluate here is

$$I = 2\int_0^{\frac{1}{2}} (1 + \rho e^{-\sqrt{4\beta y}})^{-1} dy. \tag{84}$$

We follow the identical procedure as in the previous case. First change the variable of integration to obtain

$$I = \frac{1}{\beta}\int_{-\ln\rho}^{\sqrt{2\beta}-\ln\rho} \left(\frac{x + \ln\rho}{1 + e^{-x}}\right) dx. \tag{85}$$

Assume that $\sqrt{2\beta} > \ln\rho > 0$ and write

$$\begin{aligned} I &= \frac{1}{\beta}\left(\int_{-\ln\rho}^{0} + \int_0^{\sqrt{2\beta}-\ln\rho}\right) \\ &= \frac{1}{\beta}\int_0^{\sqrt{2\beta}-\ln\rho} \frac{(x+\ln\rho)}{1+e^{-x}}\,dx + \frac{1}{\beta}\int_0^{\ln\rho}\frac{e^{-x}(\ln\rho - x)}{1+e^{-x}}\,dx \\ &= \frac{1}{2\beta}(\sqrt{2\beta} - \ln\rho)(\sqrt{2\beta} + \ln\rho) \\ &\qquad + \frac{1}{\beta}\sum_{n=1}^{\infty}(-1)^n[B_n(\sqrt{2\beta} - \ln\rho) + B_n(\ln\rho)] \\ &\qquad + \frac{\ln\rho}{\beta}\sum_{n=1}^{\infty}(-1)^n[A_n(\sqrt{2\beta} - \ln\rho) - A_n(\ln\rho)] \\ &= 1 - \frac{(\ln\rho)^2}{2\beta} - \frac{\pi^2}{6\beta} + \sqrt{2/\beta}\,\ln\,[1 + \rho e^{-\sqrt{2\beta}}] \\ &\qquad + \frac{1}{\beta}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^2}\left(\frac{1}{\rho^n} + (\rho e^{-\sqrt{2\beta}})^n\right). \end{aligned} \tag{86}$$

When $\ln\rho > \sqrt{2\beta} > 0$ we get

$$\begin{aligned} I &= \frac{1}{\beta}\int_0^{\ln\rho}\frac{e^{-x}(\ln\rho - x)}{1+e^{-x}}\,dx - \frac{1}{\beta}\int_0^{\ln\rho - \sqrt{2\beta}}\frac{e^{-x}(\ln\rho - x)}{1+e^{-x}}\,dx \\ &= \frac{1}{\beta}\sum_{n=1}^{\infty}(-1)^{n+1}\{\ln\rho A_n(\ln\rho) - B_n(\ln\rho) \\ &\qquad + \ln\rho A_n(\ln\rho - \sqrt{2\beta}) + B_n(\ln\rho - \sqrt{2\beta})\} \end{aligned} \tag{87}$$

and after collecting terms we finally obtain

$$I = \sqrt{\frac{2}{\beta}}\ln\left[1 + \frac{e^{\sqrt{2\beta}}}{\rho}\right] + \frac{1}{\beta}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^2}\left[\frac{1}{\rho^n} - \left(\frac{e^{\sqrt{2\beta}}}{\rho}\right)^n\right].$$

REFERENCES

1. George, D. A., Coll, D. C., Kay, A. R., and Bowen, R. R., "Channel Equalization for Data Transmission," The Engineering Journal (Canada), *53*, May 1970.
2. Keeler, R. J., "Construction and Evaluation of a Decision Feedback Equalizer," Rec. IEEE Int. Conf. Commun., Montreal, Canada, June 14–16, 1971.
3. Price, Robert, "Nonlinearly Feedback-Equalized PAM vs Capacity for Noisy Linear Channels," Rec. IEEE Int. Conf. Commun., Philadelphia, Pa., June 19–21, 1972.
4. Monsen, P., "Feedback Equalization for Fading Dispersive Channels," IEEE Trans. Info. Theory, *IT-17*, January 1971.
5. Doob, J. L., *Stochastic Processes*, New York: John Wiley & Sons, Inc., May 1967, pp. 159–161.
6. Berger, T., and Tufts, D. W., "Optimum Pulse Amplitude Modulation, Part: I: Transmitter-Receiver Design and Bounds from Information Theory," IEEE Trans. Info. Theory, *IT-13*, April 1967.
7. Lucky, R. W., "Decision Feedback and Faster-than-Nyquist Transmission," Paper 7.6, Abstracts 1970 Int. Symp. Info. Theory, Noordwijk, The Netherlands, June 15–19, 1970.
8. Miyakawa, H., and Harashima, H., "Capacity of Channels with Matched Transmission Technique, for Peak Transmitting Power Limitation," Nat. Conv. Rec. Inst. Elec. Commun. Eng. Japan, No. 1268, August 1969, p. 1269.